- Observational studies vs. experiments
- How to design an experiment well

- Why random sampling is important
- Which random sampling method to use 4.1
- What makes surveys so difficult
- Observational studies vs. experiments 4.2
- How to design an experiment well
- Scope of inference for sampling and experiment
- Data ethics 4.3

Sometimes It's best
just to read.

15 min ±

---

1. What is an observational study?

No treatments imposed.
- We only observe

2. What is the difference between an (explanatory) variable and a (response variable)?

outcomes

Used to predict

3. What does it mean for two variables to have an association?

- Knowing the value of one variable helps predict the other (e.g., GPA and SAT).
- *Vitamin D is associated with good health outcomes. We can predict that a person with higher D concentration will be healthier than a person with lower vitamin D concentration.*

4. If there is an association between two variables, should we conclude that there is a cause-and-effect relationship?

- *Not necessarily! In an observational study, there could be many differences between groups, not just the explanatory variable. Any of these variables could be causing the change in the response.*

5. What is confounding?

Two variables are confounded if it is impossible to determine which variable is causing a change in the response variable.

"Women who wear bikinis to the beach are less likely to have grey hair.

"Want to avoid going grey? Wear a bikini!"

6. What is an experiment? What is the primary benefit of using an experiment rather than an observational study?

7. What is a placebo?

---

8. **Define the following terms:**

- Treatment

what is done to participants

- Experimental unit/subject

What/Who is the treatment imposed on.

- Factor

When there are multiple explanatory variables.

- Level

Factors have levels

How To Buy Happiness

A video about a multi-factor experiment.

Does SAT prep improve scores ?

Check Your Understanding

SAT

Lesson 4.2: Day 1: Does SAT prep improve scores?

Suppose last year Sheldon HS offered an <u>after school</u> SAT prep class that students could volunteer to take. 44 students took the course and then took the SAT. The average SAT score for this group was 1220. The average SAT score for all students who did not take the prep class was 1050.

ABCDE

1. Is the situation described an observational study or an experiment?

2. Identify the explanatory variable and the response variable.

---

1. Is the situation described an observational study or an experiment?

Observational study, the students are not forced to take the course.

2. Identify the explanatory variable and the response variable.

Explanatory → Whether or not they took the course

Response → SAT score

d

Observational study → No treatment imposed.

Used to predict outcomes

1. Is the situation described an observational study or an experiment?

   Observational study, the students are not forced to take the course.

2. Identify the explanatory variable and the response variable.

   → Explanatory → Whether or not they took the course

   → Response → SAT score

3. Can you conclude that taking the prep course will cause a student's SAT score to increase? Why or why not?

4. Identify as many other possible variables that you can that may explain why the SAT scores are higher for those who took the prep course than for those who did not.

3. Can you conclude that taking the prep course will <u>cause</u> a student's SAT score to increase? Why or why not?

No, students who take the course might be highly motivated so they study more anyways.

4. Identify as many other possible variables that you can that may explain why the SAT scores are higher for those who took the prep course than for those who did not.

- All students in course are trying to go to college.
- They might not have afterschool jobs or be in extra curriculars.
- They have rides home from school.
- They want to do well on SAT.

---

3. Can you conclude that taking the prep course will <u>cause</u> a student's SAT score to increase? Why or why not?

No, students who take the course might be highly motivated so they study more anyways.

Need an experiment

4. Identify as many other possible variables that you can that may explain why the SAT scores are higher for those who took the prep course than for those who did not.

Confounding Variables

- All students in course are trying to go to college
- They might not have afterschool jobs or be in extra curriculars.
- They have rides home from school.
- They want to do well on SAT.

5. Design an experiment that would allow us to determine if the SAT prep causes an increase in SAT scores. Be as thorough as possible.

Experiment units

Randomly choose 100 juniors to be a part of the experiment. Randomly split them up so half of the take → the class and the others don't. After the course is over all 100 juniors take the SAT. We compare the average scores of the groups.

Treatments

Group 1
(22 students) → Treatment 1
SAT prep class

Random
assignment

ROUGHLY
EQUIVALENT

ONLY DIFFERENCE
IS TREATMENT

44 students →

Group 2
(22 students) → Treatment 2
No prep class

Compare
SAT
scores

Experiment
units

5. Design an experiment that would allow us to determine if the SAT prep causes an increase in SAT scores. Be as thorough as possible.

Treatments →

Randomly choose 100 juniors to be a part of the experiment. Randomly split them up so half of the take the class and the others don't. After the course is over all 100 juniors take the SAT. We compare the average scores of the groups.

---
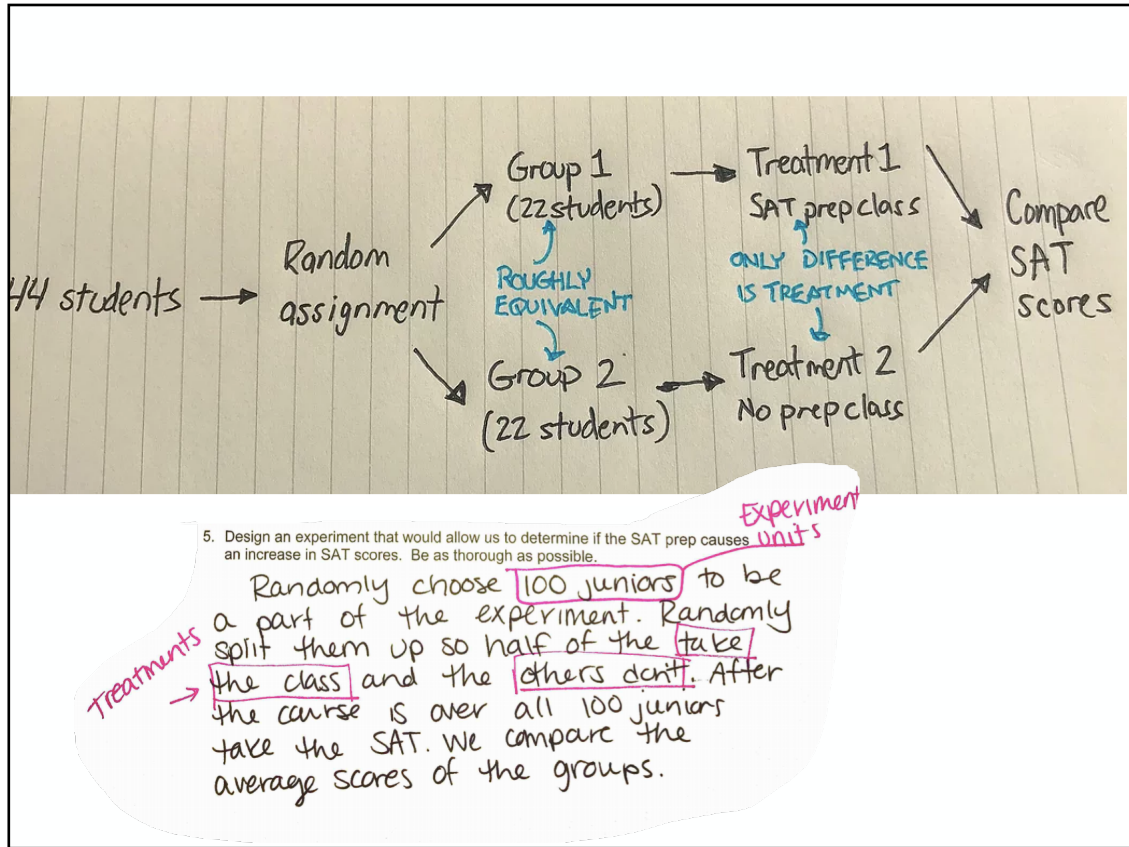
6. **Alcohol and GPA** *Confounding*

In a recent study of about 13,900 college freshman, a researcher found that the more time students spent drinking alcohol, the lower their grade point averages (GPA). Explain how confounding makes it unreasonable to conclude that spending more time drinking causes a decrease in GPA for college freshman.

6. **Alcohol and GPA** *Confounding*

In a recent study of about 13,900 college freshman, a researcher found that the more time students spent drinking alcohol, the lower their grade point averages (GPA). Explain how confounding makes it unreasonable to conclude that spending more time drinking causes a decrease in GPA for college freshman.

Possible answer

It is possible that the students who spend more time drinking are also less motivated toward academic success.

If students with lower motivation have lower GPA's, it could be that the students' motivation level caused the GPAs to go down, not the time drinking.

7. **The best test scores** ---- *Vocabulary of experiments*

Several AP® Statistics students wondered whether caffeine could improve test scores. They randomly assigned 30 student volunteers to either drink regular coffee or decaffeinated coffee the morning of the students' next test. At the end of the experiment, they recorded test scores for each student volunteer. Identify the treatments and the experimental units in this experiment.

7. The best test scores ---- *Vocabulary of experiments*

Several AP® Statistics students wondered whether caffeine could improve test scores. They randomly assigned 30 student volunteers to either drink regular coffee or decaffeinated coffee the morning of the students' next test. At the end of the experiment, they recorded test scores for each student volunteer. Identify the treatments and the experimental units in this experiment.

This experiment compares two treatments:
(1) regular coffee and (2) decaf coffee

The experimental units are the 30 student volunteers.

8. Growing the best tomatoes ---- *Experiments with multiple explanatory variables*

Does adding fertilizer affect the productivity of tomato plants? How about the amount of water given to the plants? To answer these questions, a gardener plants 24 similar tomato plants in identical pots in his greenhouse. He will add fertilizer to the soil in half the pots. Also, he will water 8 of the plants with 0.5 gallon of water per day, 8 of the plants with 1 gallon of water per day, and the remaining 8 plants with 1.5 gallons of water per day. At the end of 3 months, he will record the total weight of tomatoes produced by each plant.

8. Growing the best tomatoes ---- *Experiments with multiple explanatory variables*

Does adding fertilizer affect the productivity of tomato plants? How about the amount of water given to the plants? To answer these questions, a gardener plants 24 similar tomato plants in identical pots in his greenhouse. He will add fertilizer to the soil in half the pots. Also, he will water 8 of the plants with 0.5 gallon of water per day, 8 of the plants with 1 gallon of water per day, and the remaining 8 plants with 1.5 gallons of water per day. At the end of 3 months, he will record the total weight of tomatoes produced by each plant.

1. List the factors in this experiment and the number of levels for each factor.

2. If the researchers used every possible combination to form the treatments, how many treatments were included in the experiment?

Factors

1. Whether the fertilizer is applied (2 levels) and the amount of water (3 levels)

2. $2 \times 3 = 6$ different treatments.

3. Treatments
   (1) fertilizer, 0.5 gallon
   (2) fertilizer, 1 gallon
   (3)   etc

LCQ

Assignment:

4.2......43, 45, 47, 49, 51, 53